Modern Practice of Attribution on Administrative Claims Data

Paul Rathouz, PhD

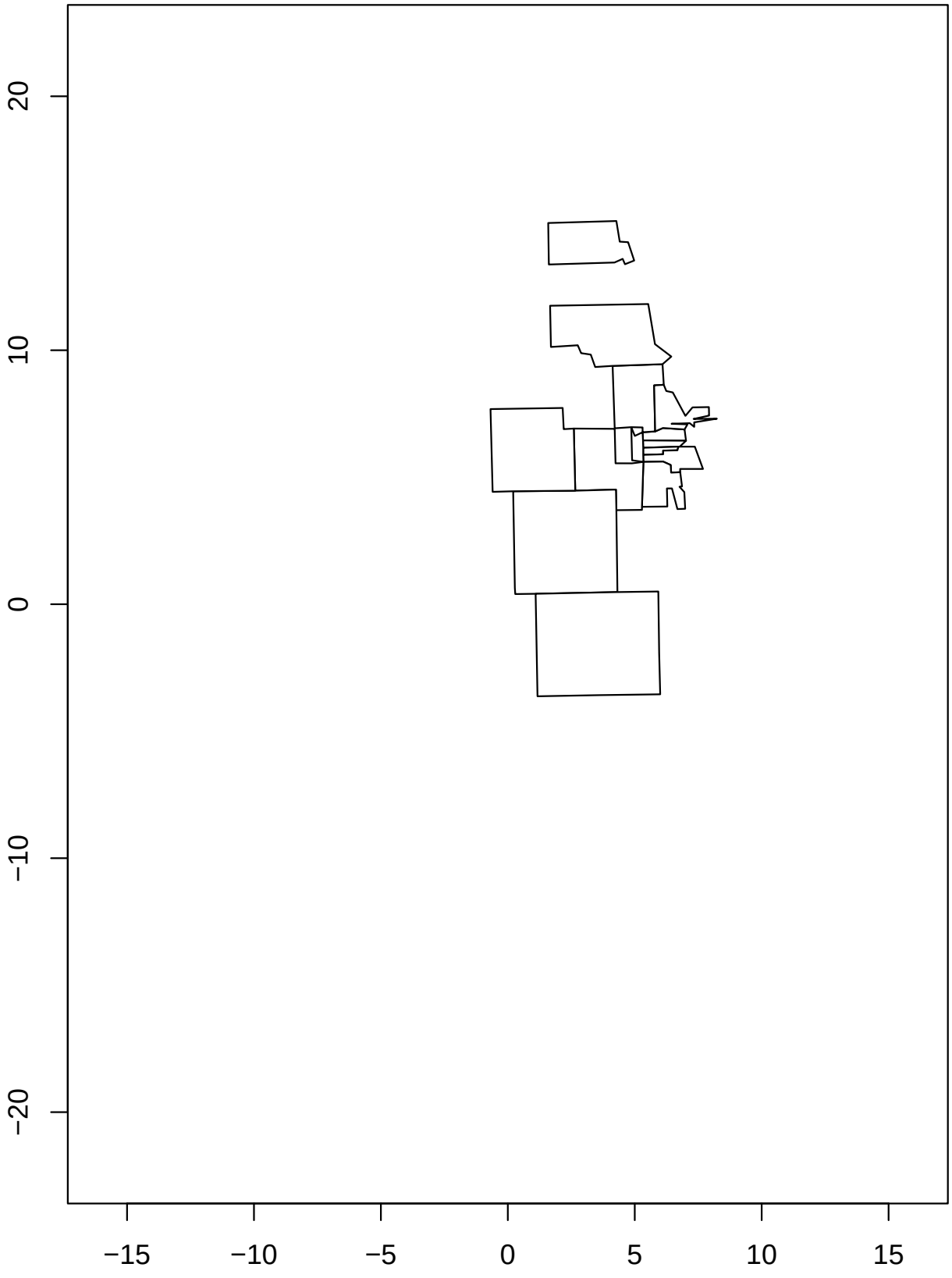
prathouz@uchicago.edu

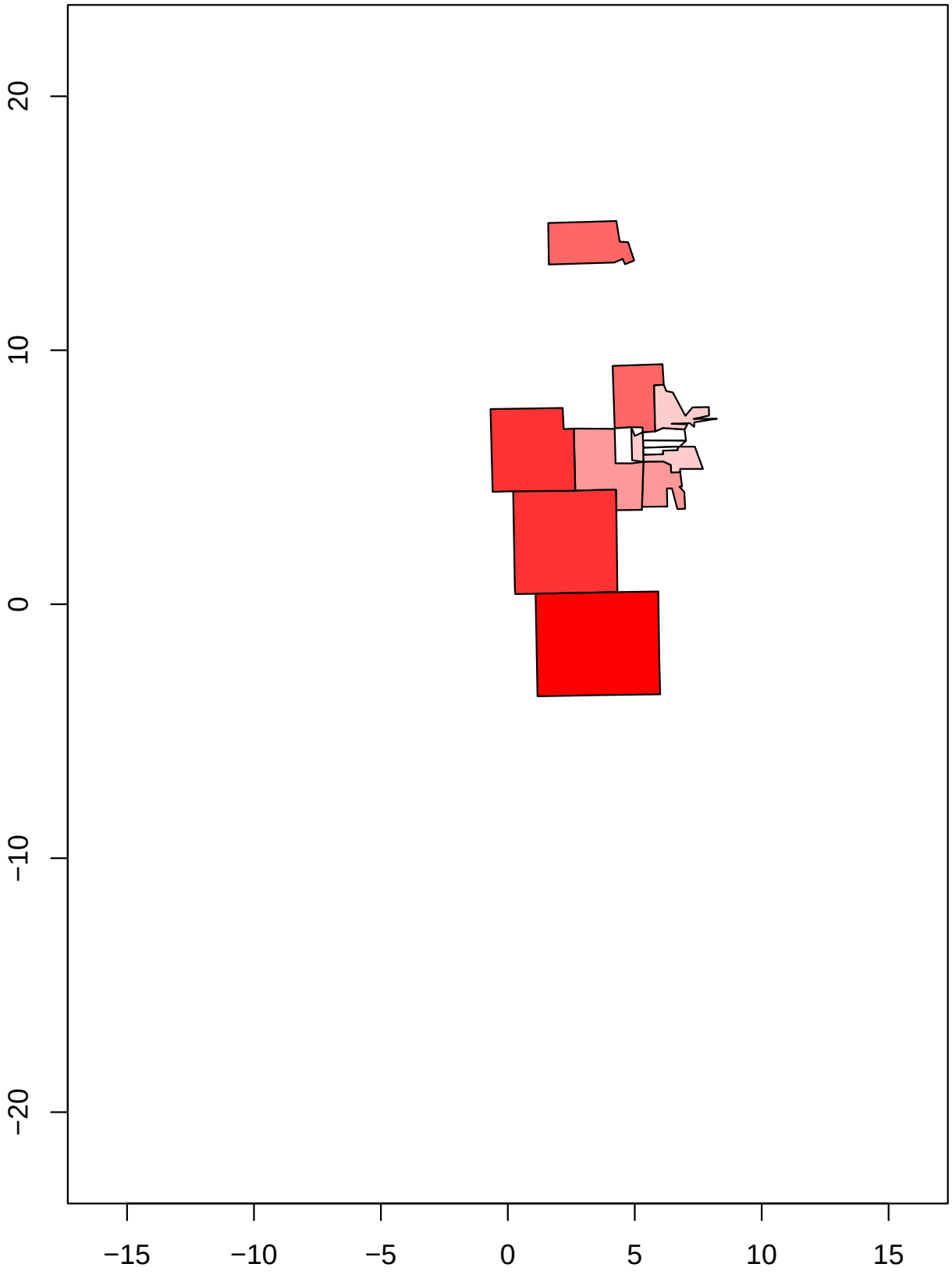
Department of Health

C [illegible] o A [illegible] M d [illegible] d [illegible]

IDPA 2 people on Medicaid in Chicago

Medicaid (IDPA) claims data


20
10
0
−10
−20
−15
−10
−5
0
5
10
15


20
10
0
−10
−20
−15
−10
−5
0
5
10
15

Creating Air Pollution Exposure Estimates

Ozone: 11 monitors

- **Hourly ozone**
- **Max-8-hr average ozone**
- **Ordinary spatial Kriging assuming isotropy**
- **Matern covariance function**
- **Predicted values of ozone (and se) at each ZIP centroid**
- **Spatially resolved exposure measure!**

PM10: 18 monitors

- **Observations spaced 6 days apart**
- **Monitor and day effect ANOVA model to predict PM10 for any given monitor on any given day**
- **Daily average based on the fitted values**

Weather: Average of O'Hare and Midway:

Prediction of BA on BA

Why use BA (Albuterol)?

- often-used non-steroid bronchodilators
- specific to asthma
- running prescriptions, quick use
- large counts

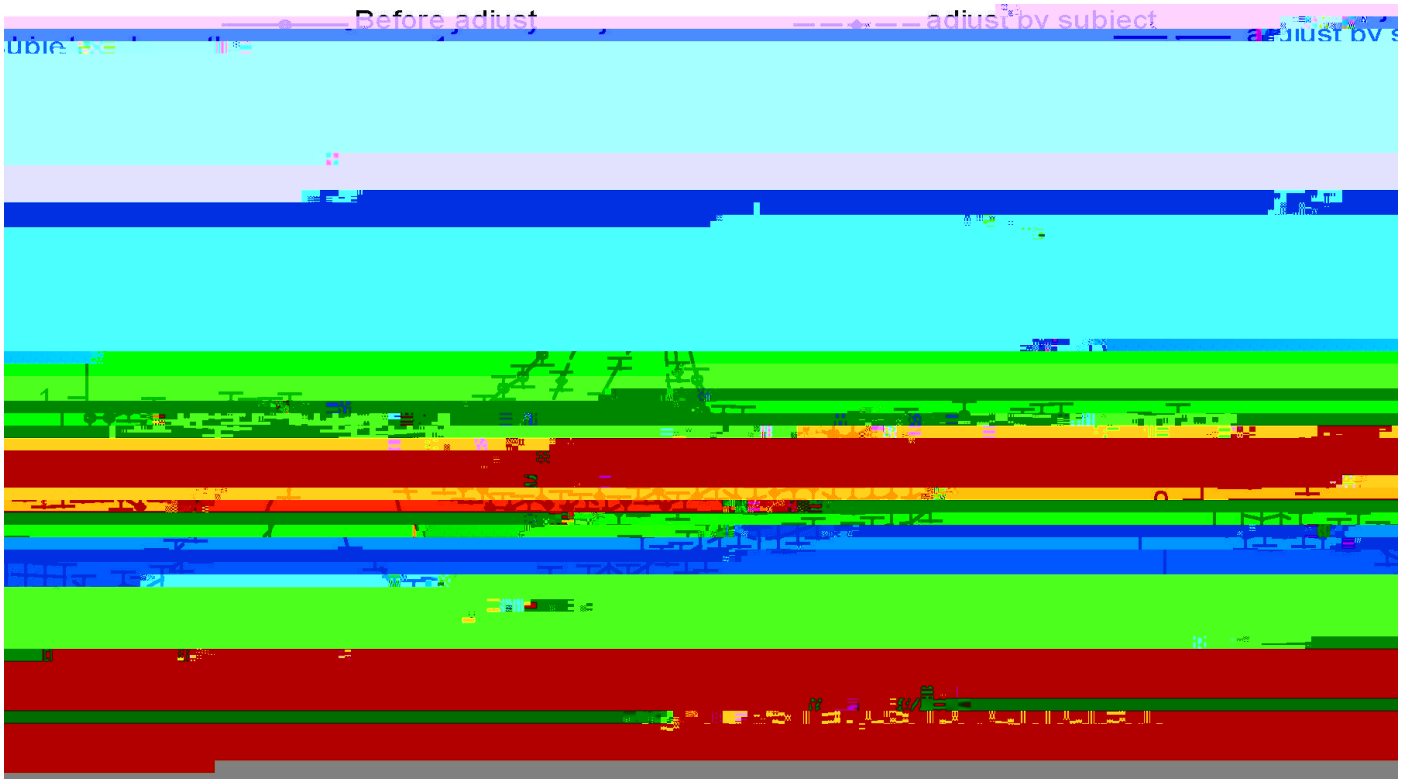
Q: Is using BA prescriptions as a marker for asthma outcome reasonable?

- Examine association between BA and hospital admits / ED visits (more traditional outcomes), lagged by one or more days

Odds ratio (OR)

- crude OR
- subject-adjusted OR
- subject- and time-adjusted OR

A oc· ·on o n BA nd
D o p Ad ·

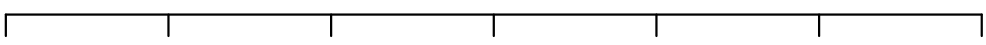


(Note: positive lag BA before ED)

Problems with crude OR:

. . . Artifactually

M pp n d co d y
cod



observed, (over space and time)

disease risk within ZIP code (

p d c d disease risk (given data)

po n y po n

ay od . ˙ ˙ on ˙n d˙- ˙ ˙-

description of disease incidence

. . . spatial interpolation

! disease risk ˙ pp˙n

place spatial epidemiologic studies in context by displaying background risk

interpret (variation in) disease risk at po˙n versus ˙ level

- **smooth variation across areal units**
- **account for (di erences in) sampling variability across areal units**
- **borrow information across units**

understand areal data aggregation

improve statistical e ciency in spatial regression models

Approximate Modeling Assumption

R is constant at location

Then ...

$$Y_i \sim \text{Poisson}\left(\text{person-time} \cdot e^{\bar{R}_i}\right)$$

where R_i = average of R over ZIP i

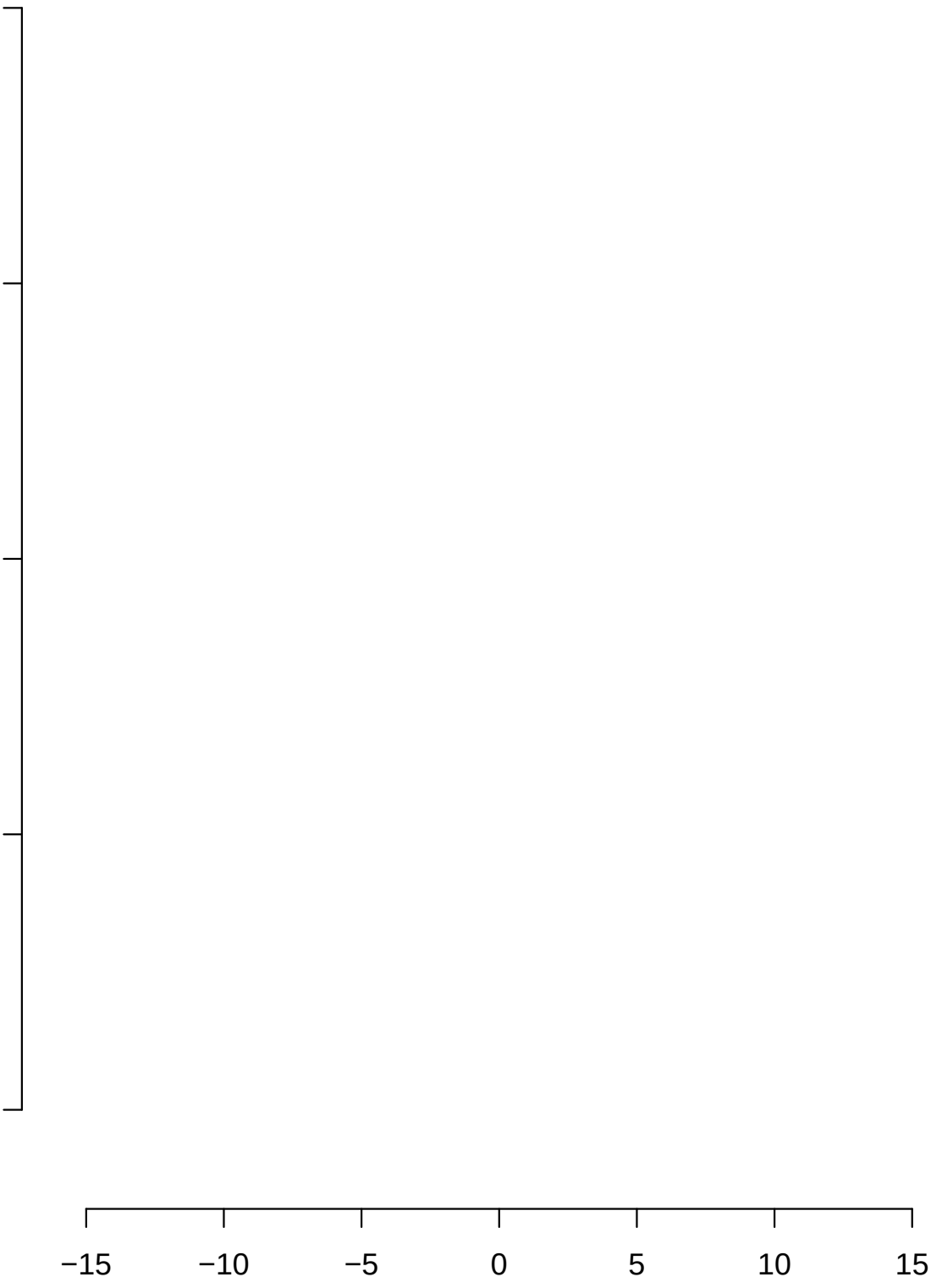
Converts point process problem in R to simple problem in R_i

An estimate of R_i is

$$\widehat{R}_i = \log\left(\frac{\#\text{ events}}{\text{person-time}}\right)$$

the observed log risk in ZIP i

$\widehat{R}_i$'s are used to predict R at every


−15
−10
−5
0
5
10
15

M o A p c of o c

1. p c ozon od n c o n
   new statistical space-time models
   exploit physical models such as CMAQ
   improve air pollution exposure measurement
2. Validation of Beta-agonist as Asthma Outcome
3. Mapping asthma outcomes aggregated by ZIP-code
4. A d y of co
5. p on d y of co

C zon A
Co n o

Vanja Dukic (Health Studies)

Dana Draghicescu (CISES, Research Associate)

John Frederick (Geophysical Sciences)

Edward Naureckas (Medicine)

Alexis Zubrow (CISES, Programmer/Analyst)

Xiaoming Bao (Statistics, MS Student)

Lon [illegible] d[illegible] An[illegible]y[illegible] of A[illegible] [illegible] co

D[illegible]

Y_{ijt} I event, person j, ZIP i, day t

i ZIP; j person in ZIP; t day

ij (if no[illegible] d) location

[illegible] d od[illegible]

$$Y_{ijt} \quad e^{R(x_{i})+S_{i} + \prime z_{i t}}$$

where:

R is baseline o[illegible] [illegible] at

S_{ijk} all unobserved factors for person i j in time \window" k

z_{ijt} covariates for person i j, day t

Co[illegible]f o[illegible]nd[illegible]n con o[illegible]nd[illegible]c[illegible]ncy

$\widehat{\beta}$ adjusted for spatially and within-person slowly-varying factors

$\widehat{\beta}$ has improved statistical e ciency

M[illegible] zrma

Lon [illegible] d[illegible] [illegible] Da [illegible] An [illegible] y [illegible]

[illegible] [illegible] n [illegible] y Mod [illegible] [illegible]

Ad[illegible]

Odds ratios for Risk of BA Prescription Fill

Conditional Logit Model

	Odds ratio	St. Err.	Z
ozone (20ppb)	1.005	.007	0.81
pm10 (15mg/m3)	.999	.004	-0.22
log(pollen) (std)	1.009	.004	2.11
temp (5F)	.984	.003	-5.20
rel hum (10%)	.997	.003	-0.95

Average Odds Ratios across 51 ZIP Codes

	Odds ratio	St. Err.	Z
ozone (20ppb)	1.032	.022	1.47